Research on

Analog Statistical Function Discriminators Utilizing Ferrielectric Elements

Prepared for

National Aeronautics and Space Administration
Electronic Research Center
575 Technology Square
Cambridge, Massachusetts

Interim Report
October, 1966

Prepared by

Electrocristal Corporation, Inc.
2014 Taylor Street, N. E.
Washington, D. C.

Under
Contract No. NAS 12-136

TABLE OF CONTENTS

LIST OF TABLES AND ILLUSTRATIONS

Interim Report
Contract No. 12-136
October, 1966

I. RESEARCH INVESTIGATIONS

A. Program Objectives and Major Considerations.

This Interim Report has been prepared for NASA, Electronic Research Center, covering the four-month investigation conducted under Contract 12-136 during the period June, July, August, and September, 1966.

The fundamental aims of this study were directed toward the development of an analog statistical switch capable of processing information for adaptive control devices. This statistical switch may be employed in a combined logic-memory system. The concept conceived and adopted for the processing of information transforms signals, images, or events (briefly referred to as a "pattern") to be recognized into more than one discriminant function, each of which represents a distribution of a particular property of the pattern to be characterized. Accordingly, various property distributions would be processed from a particular pattern--for example, a Poisson distribution, which is well suited to describe properties of a signal or pattern such as the one present in an electrical signal known as zero crossings. Additionally, perhaps a Raleigh distribution could represent the distribution of another property of the pattern, such as modulation intensity properties and so on. Each one of these property functions would be a composite (not elementary) distribution function, which could be electronically processed, and could be used to find or derive a generalized discriminant function. A set of such generalized discriminant functions, each one derived from a particular property of the pattern, could then be used for recognizing a particular pattern presented to the machine. In a further step the mutual probabilities of these probability functions can also be utilized in the recognition process. It is assumed that the various property distribution and discriminant functions are characterized by a fairly small number of representative samples, and require first of all a multi-level storage system, which can preferably be simultaneously interrogated, and can also, in a later development, be trained.

The theoretical portion of the study undertaken may be broken down into three areas:

1. Characteristic features of various representative distribution

functions were established and tabulated.

2. Mathematical considerations were given to the dividing of arbitrary distribution functions into the sum and product of simple distribution functions.

3. An analysis was made of simple distribution functions and/or composite distribution functions, which are best suited to represent a particular property of the pattern to be recognized.

4. Consideration was given to a system in which a set of property discriminant functions are used for recognizing a particular pattern.

In order to understand the theoretical objectives listed above, certain experimental work was begun to demonstrate an electronic device capable of processing density distribution functions and storing and transforming up to sixteen samples of arbitrary functions in a mock logic-memory system.

B. Theoretical Studies.

Three tables of probability (density) functions have been compiled, and their characteristic features have been tabulated. (Hereafter these functions will be referred to simply as "distribution functions.")

1. Table I contains the more significant distribution functions, each of which can be electronically processed.

2. Table II contains discrete distribution functions. Of those enumerated, the Poisson is the only one for which we have found an electronic method of processing.

3. Table III contains those distribution functions which Medgyessy has found suitable for superposition and which he has tested mathematically.

The representation of a composite distribution function in terms of elementary distribution functions is a sophisitcated mathematical task due to the fact that a particular wave form, periodic or non-periodic, must satisfy some mathematical criteria in order to be suitable for such a decomposition process. However, Medgyessy gives in his book the mathematical criteria for those elementary distribution functions which are suitable for the representation of composite distribution functions by superpositions. According to his findings, there are about eight distribution functions (see Table III) which can be used. (The characterization of those distribution functions already given in Tables I and II are not repeated in Table III.)

#	Distribution	Probability Density Function $\phi(x)$	$\Phi(x)$ Distribution Function	$E\{x\}$ Mean Value	$Var\{x\}$ Variance	υ Skewness	Remarks
1	Gaussian	$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ $(-\infty < x < \infty)$	$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$	0	1	0	$\operatorname{erf} x = 2\,\Phi_x(x/\sqrt{2}) - 1$
2	Rectangular or Uniform Distribution	$\begin{cases} \frac{1}{2\alpha} & (\lvert x-\xi\rvert < \alpha) \\ 0 & (\lvert x-\xi\rvert > \alpha) \end{cases}$	$\begin{matrix} 0 & (x \le \xi-\alpha) \\ \frac{1}{2\alpha}(x-\xi+\alpha) & (\xi-\alpha \le x \le \xi+\alpha) \\ 1 & (x \ge \xi+\alpha) \end{matrix}$	ξ	$\frac{\alpha^2}{3}$	0	x is uniformly distributed over the interval
3	Raleigh	$\begin{cases} 0 & (x<0) \\ \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} & (x \ge 0) \end{cases}$	$1 - e^{-x^2/2\sigma^2}$	$\sqrt{\frac{\pi}{2}}\,\sigma$	$0.43\sigma^2$		The envelope of a band of normally distributed noise approaches a true Raleigh distribution as the bandwidth approaches zero.
4	Exponential	$\begin{cases} \frac{1}{\beta} e^{-\left(\frac{x-\xi}{\beta}\right)} & (x>0) \\ 0 & (x<0) \end{cases}$	$\begin{matrix} e^{-\frac{\lvert x-\xi\rvert}{\beta}} & (x \le \xi) \\ 1 - e^{-\frac{\lvert x-\xi\rvert}{\beta}} & (x \ge \xi) \end{matrix}$	$\xi + \beta$	β^2	2	Useful when the random variable is the length of an interval without the occurrence of a particular event where the average rate of occurrence of the event is β.
5	Causal or Dirac Delta	$\delta(x-\xi) = \begin{cases} 0 & (x \neq \xi) \\ \infty & (x = \xi) \end{cases}$	$U(x-\xi)$	ξ	0	Not defined	x is almost always equal to ξ. Note that the rectangular and Cauchy approximate a causal distribution as $\alpha \to 0$ or $\beta \to 0$
6	Lorentzian	$\frac{\alpha}{\beta^2 + x^2}$	$\frac{\alpha}{\beta} \arctan \frac{x}{\beta}$	0	1	0	An LC resonance curve describes a Lorentzian distribution.

TABLE I.
Continuous Probability Distributions

#	Distribution	Point Probabilities	Restrictions on Parameters	$E\{x\}$ Mean Value	$Var\{x\}$ Variance	ν Skewness	Remarks
1	Poisson	$e^{-\xi}\dfrac{\xi^x}{x!}$	$0<\xi<\infty$	ξ	$\xi^{-\frac{1}{2}}$	ξ	Domain: $x = 0, 1, 2, \ldots \infty$; $\xi > 0$
2	Geometric	$\nu(1-\nu)^x$	$0 \le \nu \le 1$	$\dfrac{1-\nu}{\nu}$	$\dfrac{1-\nu}{\nu^2}$	$\dfrac{2-\nu}{\sqrt{1-\nu}}$	Domain: $x = 0, 1, 2, \cdots \infty$
3	Binomial	$\binom{n}{x}\nu^x(1-\nu)^{n-x}$	$0 \le \nu \le 1$	$n\nu$	$n\nu q$	$\dfrac{q-\nu}{\sqrt{n\nu q}}$	$q = 1-\nu$ Domain: $x = 0, 1, 2, \ldots, n$
4	Polya's or negative binomial	$p(x)=\left(\dfrac{\xi}{1+\beta\xi}\right)^x \dfrac{1(1+\beta)\ldots[1+(x-1)\beta]}{}p(0)$ $p(0)=(1+\beta\xi)^{-\frac{1}{\beta}}$	$(\xi>0, \beta>0)$ $(x\ge 0)$	ξ	$\xi(1+\beta\xi)$	See (2)	1) Polya's distribution reduces to the Poisson distribution for $\beta = 0$, and to the Geometric distribution for $\beta = 1$ 2) $\nu = \dfrac{(1+\beta\xi)+\xi/n}{\sqrt{\xi(1+\beta\xi)}}$ $(n \ge 0)$

TABLE II.
Discrete Probability Distributions

#	Distribution	Probability Density Function $\phi(x)$	$\Phi(x)$ Distribution Function	$E\{x\}$ Mean Value	$Var\{x\}$ Variance	υ Skewness	Remarks
1	Cauchy	$\frac{1}{\pi\alpha}\frac{1}{1+\left(\frac{x-\xi}{\alpha}\right)^2}$	$\frac{1}{2}+\frac{1}{\pi}$ arctan $\frac{x-\xi}{\alpha}$	$E\{x\}$ and $Var\{x\}$ do not exist; the Cauchy principal value of $E\{x\}$ is ξ.		Not defined	Distribution of $x=\xi+\alpha \tan y$ if y is uniformly distributed between $y=-\pi/2$ and $y=\pi/2$ (rectangular distribution). Cauchy's distribution is symmetric about $x=\xi$. Half width and interquartile range are both equal to 2α.
2	Beta	$0 \quad (x\leq 0, x\geq 1)$ $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ $(0<x<1)$ $(\alpha>0, \beta>0)$	$0 \quad (x\leq 0)$ $I_x(\alpha,\beta) \quad (0\leq x\leq 1)$ $1 \quad (x\geq 1)$	$\frac{\alpha}{\alpha+\beta}$	See (1)	$\frac{2(\alpha-\beta)}{(\alpha+\beta+2)}$	1) $Var\{x\}=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ 2) $I_x(\alpha,\beta)$ is the complete beta function ratio; unique mode $(\alpha-1)/(\alpha+\beta-2)$ for $\alpha>1$, $\beta>1$.
3	Gamma	$0 \quad (x\leq 0)$ $\frac{1}{\beta^\alpha\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$ $(x>0)$ $(\alpha>0, \beta>0)$	$0 \quad (x\leq 0)$ $\frac{1}{\Gamma(\alpha)}\Gamma_{x/\beta}(\alpha)$ $(x\geq 0)$	$\alpha\beta$	$\alpha\beta^2$	$2/\sqrt{\alpha\beta}$	$\Gamma_x(\alpha)$ is the incomplete gamma function.
4	Logarithmic Normal	$0 \quad (x\leq\alpha)$ $\frac{1}{(x-\alpha)\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{\log_e(x-\alpha)-u}{\sigma}\right]^2}$ $(x>\alpha)$	$\int\phi(u)du$ $u=\frac{\log x-\log\xi}{\sigma}$ $(0<x<\infty)$	See (1)	σ^2	$\sigma \log x$	1) $E\{x\}=\xi\times 10^{\frac{\sigma^2}{2\log e}}$

Note: The Gaussian, Poisson, Binomial, and Polya distributions are also suitable for superposition.

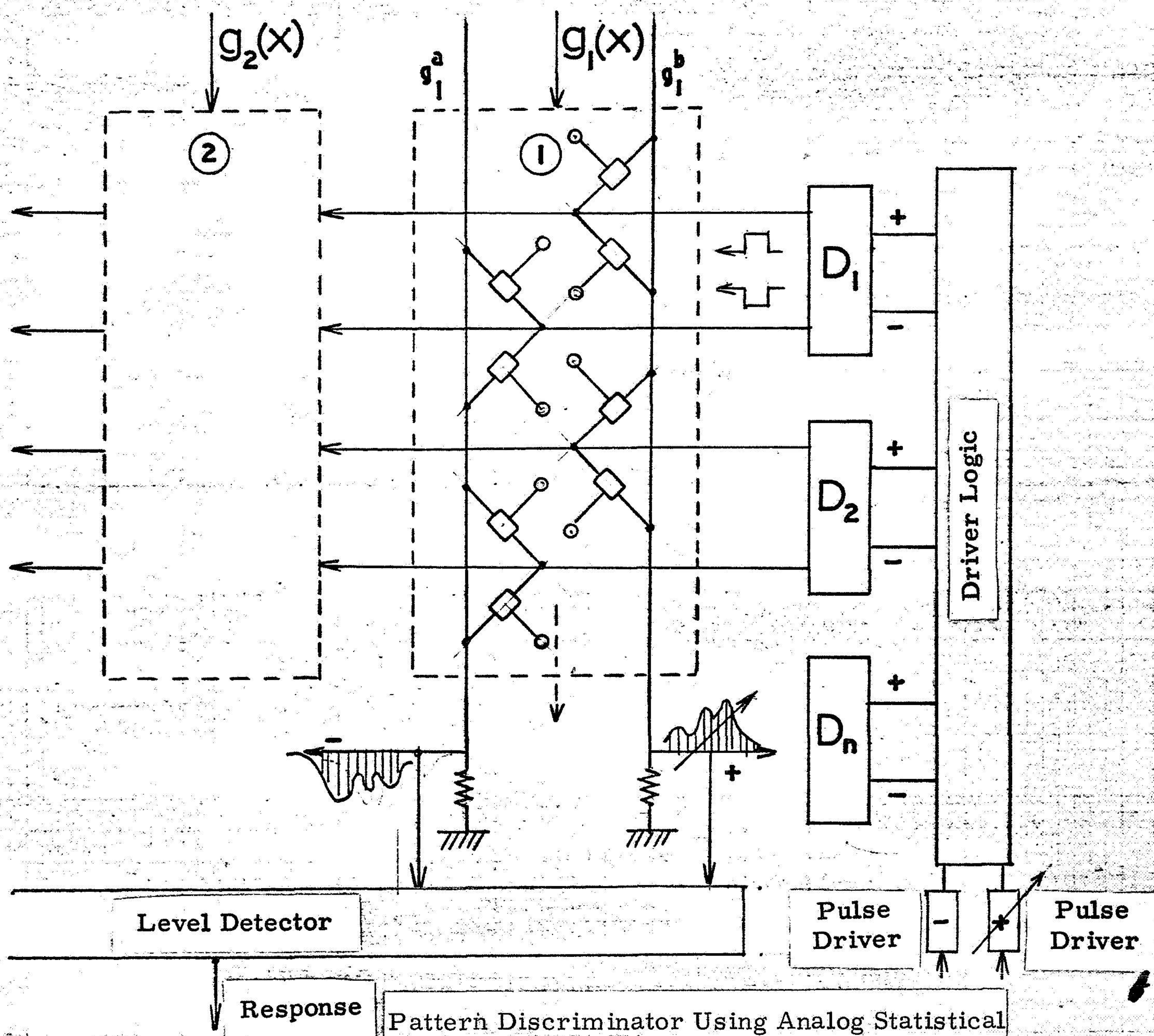
TABLE III.

Continuous Probability Distributions Suitable for Superpositions

An analog statistical switch may be defined as a switch which responds to a particular distribution of analog signals. Such a switch may sometimes be regarded as a special filter. It may also be looked upon as a pattern discriminator.

As a conceptual example of a pattern discriminator using the analog statistical switches just described, or, more accurately, an analog associative memory, Figure 1 is included. The property distribution processors $\phi_i(x)$ give the distributions of the various properties of a particular pattern, such as frequency spectrum, zero crossings, density distribution of peaks with amplitude, etc. The output of each distribution processor feeds into a set of statistical switches gj $\left[\phi_i(x)\right]$ representing typical distributions within a particular (i) class of properties. The outputs of the various groups of statistical switches are discriminated in a flag, and the outputs of the flag are further processed in the response device (shown schematically). More details are given later in this report.

The focal point of this study was to determine the conditions which must be met by an electronic processor to extract the distribution of a particular property such as frequency spectrum, zero crossings, etc., from an electrical signal. From studying the mathematical properties of a particular type of distribution, it becomes clear from both the extraction of sets of samples and the organization of these sets into a distribution , which properties can be best characterized by its use. For example, the Poisson distribution is well suited to describe a process such as the one present in an electrical signal known as zero crossings, which would appropriately characterize the conture lines of a picture or simply give a line drawing of a picture. In contrast, a pure, normal distribution describes events which are not interrelated and give amplitude information useful for filling in the line drawing characterized by the Poisson distribution. According to this concept for characterizing the various properties, appropriate distribution functions are used. This characterization of the various distribution functions will later permit the specification of various electronic devices which are suitable for the extraction of some particular features from an electrical signal. Table IV gives examples of properties the distributions of which may be most aptly described by a particular distribution function. Although this work actually constitutes only a beginning step toward the characterization of electrical signals by some of their composite (not elementary) distribution functions, it served, however, the purpose of offering a more realistic approach to the design requirements of a multi-level logic-memory, which, as a result of the discussions conducted with Dr. David Van Meter on September 29th, may now be appropriately referred to as an "analog associative memory."



Pattern Discriminator Using Analog Statistical Switches

Figure 1

Normal Distribution	The statistics of peaks of a random signal over the amplitude range. It is suitable to characterize the amplitude features of a signal. Composite normal distributions would have the same features.
Rayleigh Distribution	The statistics generated by the linear (envelope) detection of a narrow band of Gaussian noise. It is characterized by the distribution of the envelope of the peaks and represents modulation intensity properties.
Poisson Distribution	The statistics of zero crossings of band-limited white noise. It is suitable to characterize the frequency feature of a signal without obtaining the intensity of the signal.
Frequency Distribution	The statistics of one countable number of possible values of a random signal such as frequency of occurrence. It is suitable to characterize spectral density.

TABLE IV
Examples for the Characterization of Some Properties of Electrical Signals by Probability Distributions

C. Analog Associative Memory.

Based on the theoretical investigation undertaken, and our discussions with NASA personnel, the need for work on an analog associative memory has developed as one of the primary areas for investigation under this contract. An analog associative memory would permit one to compile and interrogate complex information by handling whole functions in the computational process. The ferri-electric memory-logic system advanced by Electrocristal will permit one to handle multi-level storage features for any one representative property vector.

It should be noted that before the design of an analog associative memory can be undertaken, a number of questions need to be answered, some of them of a theoretical, others of a practical, nature. These include:

1. In what form will an analog signal be presented to the associative memory?

2. Will the analog signal presented be phase sensitive?

3. Will the interrogating signals have some deviation from the distribution functions stored in the analog associative memory?

4. If there should be a deviation from the stored discriminant function, what would be the nature of this deviation? Is it a random deviation? And, regardless of the nature of the deviation, what type of flexibility could be built into the associative memory so as to take care of such fluctuation?

5. Is a normalization of the input signal required?

6. How should the structure of the matrix be organized?

7. What type of sensing flag should be used to detect a response?

8. Should the interrogation of the memory matrix be parallel or sequential, or should it be parallel and sequential?

Although the advantages of an associative memory are well known to those who are actively interested in seeking new communication techniques and software, little or no work has been done on analog associative memories. Some of the more notable problems associated with developing such a memory may be listed as follows:

1. Multi-level storage

2. Simultaneous interrogation of a large matrix

3. Signal to noise ratio for equality matches, etc.

4. Statistical deviations between interrogating and stored functions

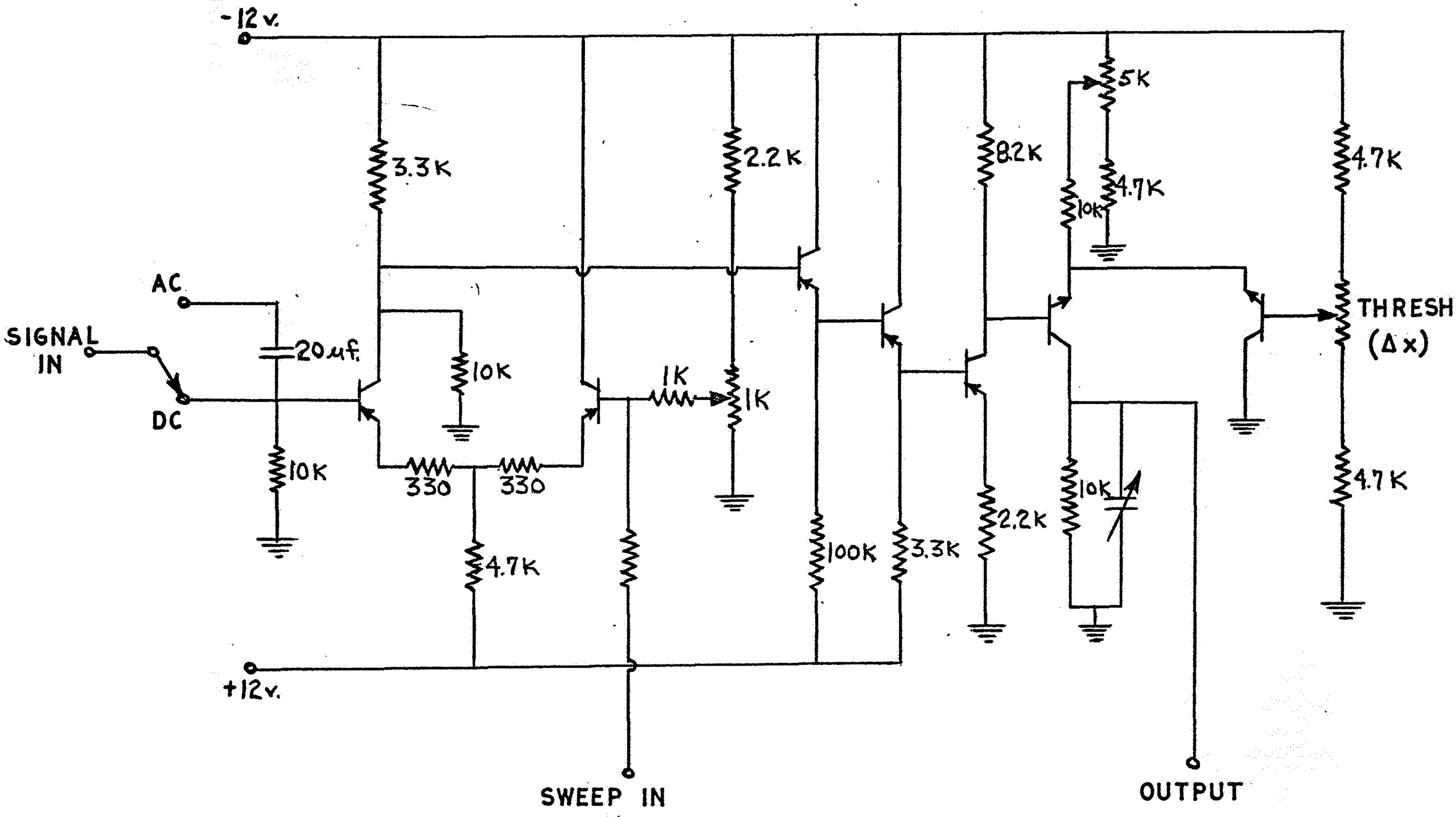
D. Experimental Investigations.

An experimental model for processing density distribution functions from various types of input signals was built and tested. A circuit diagram of this model is attached. (See Figure 2.) The basic principle of operation of this circuit is, briefly, as follows: A trigger circuit is operated by the peaks of the signal as the trigger level is slowly swept across the whole signal amplitude. On the output, the number of peaks can be integrated, and the integrated signal can be displayed on a scope. The sweep frequency must be slow in order to obtain a sufficient statistical average. In the simplest sense, this type of circuit gives, for example, a Gaussian distribution function if the input receives a random noise signal. This circuit is a typical distribution function processor.

It has become obvious that the speed of such a processor is too slow, and for this reason it is desirable to develop a higher speed device. The principle of such a high speed device can be outlined as follows:

The signal would be fed into a multi-level detector. On the output of each level, the pertinent peak densities can be integrated. Since such a processor would not be a sequential, but rather a parallel operating device, the speed could be increased to a suitable level. The development of such a device is at this moment deferred to a later date, since the most important task in this research period is to define and develop a statistical switch incorporating the various novel features already described.

A basic building block approach was followed in order to establish future adaptive control devices which include learning features. For this purpose, a logic memory storage system capable of handling sixteen samples is planned. The objective of this experimental investigation is to provide hardware which can expand and compress a particular distribution function on a controlled experimental basis. In other words, it would become possible to generate families of similarity (affine) functions, and these functions would represent a variable pitch internal model with which an incoming signal can be correlated. In the final solution, feedback and synchronizing features will also be incorporated.



DENSITY DISTRIBUTION PROCESSOR
Figure 2

The experimental work described above is designed to implement this flexible internal model or flexible word. The build-up of this experimental model requires the development of the following items:

1. Driving cards.

2. A sequential flip-flop chain, the switching speed of which is variable.

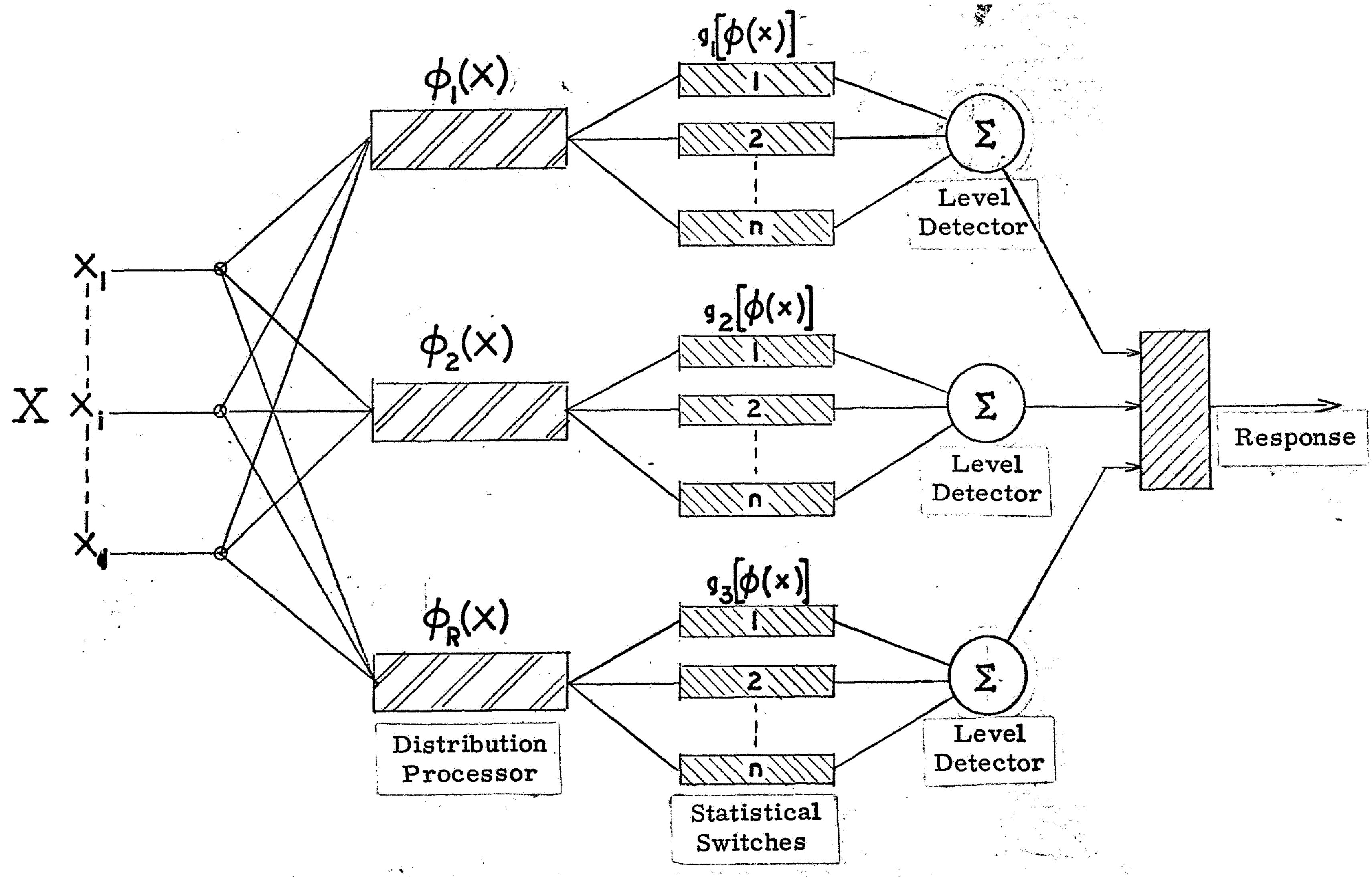
3. A sixteen-bit Transpolarizer memory unit, each bit of which contains two Transpolarizers, one of which produces the flexible internal model, and the second of which contains the signal to be interrogated.

4. Various auxiliary circuitries.

Figure 3 represents a preliminary concept of an analog associative discriminator denoted by $g_1 [\phi_1(x)]$, $g_2 [\phi_1(x)] \ldots g_n [\phi_1(x)]$, having a flexible fluctuating internal distribution function represented by discrete samples of the just mentioned distribution function and by the settings of the Transpolarizers in Column 1 b, while the incoming signal can be simulated by the settings of the Transpolarizers in Column 1 a. The fluctuations may represent any sort of linear or nonlinear transformations and afford the producing of desired similarity distribution functions in a clustered manner around a typical distribution. For the time being, an arbitrary number of sixteen Transpolarizers will be used for this experimental model. The Transpolarizers in each column can be set to arbitrary levels by the application of a potential to the control points of the Transpolarizers. Each Transpolarizer column can be sequentially driven by the drivers D 1, D 2, ...D_n. The drivers are driven by a driver logic which generates sixteen sequential outputs as a clock pulse input drives the logic. This driver logic is constructed in such a manner that after a single sequence of driving pulses has switched the sixteen Transpolarizers through a complete cycle, an additional seventeenth driving board is activated. This driving board sends out an opposite polarity restoring pulse and switches all sixteen Transpolarizers back to their original stand-by condition.

This very same driving system is applied to the flexible internal stored function as well as to the Transpolarizers which simulate a particular distribution function of an incoming signal.

As Figure 3 shows, the switching is for the time being so set that opposite polarity outputs are obtained from the two Transpolarizer columns a and b. This way it becomes possible to subtract the incoming distribution from the model distribution, and to obtain a difference distribution for further



Experimental Analog Associative Discriminator

Figure 3

use. It also will become possible to drive the internal model with a variable frequency and as a consequence a cluster of similarity functions can be produced for comparison purposes. The transformation of the various similarity functions can also be linearly or nonlinearly accomplished.

Although in the first breadboarded models discrete Transpolarizers will be used, in future devices a memory element representing a whole multi-level analog function can be fabricated in a flat pack and would be comparable to a whole memory word in an ordinary computer. Large matrixes of such structures will become feasible later where, in a cross point, so to speak, we may have a whole analog function instead of a binary "1" or "0."

It is not considered necessary to make a detailed comparison between a simple binary matrix and a matrix having whole analog functions stored and organized for interrogation. It is apparent that while in a binary representation the possible number of words is 2^n, the possible number of words in a multi-level representation having "n" levels becomes n^n.

II. INTERPRETATION OF RESULTS

In summary, it can be stated that the theoretical investigations show interesting possibilities for:

1. A system utilizing memory-logic elements

2. Flexible stored functions

3. Analog associative memories.

Based on the preliminary studies, a modest laboratory mock-up is being assembled. This apparatus will be employed in an experimental investigation to further document the feasibility of handling whole analog functions represented by samples in a memory process. This work has further shown that a key portion of the entire concept will involve an analog associative memory, and thus the research efforts for the coming research periods shall be concentrated on such an analog associative memory.

III. RECOMMENDATIONS FOR FUTURE ACTION

In the first phase of Contract No. 12-136, Electrocristal undertook the development of analog statistical function discriminator concepts which could be employed in adaptive control and learning systems. This work has pointed out the need for additional investigations in that portion of an adaptive computing system which may be defined as an analog associative memory. We therefore propose that a second phase of the program be initiated, directed specifically towards the establishment of concepts and demonstration laboratory hardware which will prove the feasibility of an analog associative memory employing ferrielectric storage elements.